Evaluating crowding in individual train cars using a dynamic transit assignment model

Soumela Peftitsi · Erik Jenelius · Oded Cats

Received: date / Accepted: date

Abstract Crowding is one of the major issues of public transport systems and has many negative effects for both users and operators. Passengers can be highly unevenly distributed between individual cars of a train even when the total passenger load exceeds the practical capacity.

Transit assignment models are widely used for describing and evaluating crowding in the vehicle. However, these models usually do not capture how passengers are distributed across the vehicle. To this end, this study introduced the capability to analyze the effective capacity utilization of the train, by considering passengers’ distribution among individual train cars into an agent-based simulation model.

The developed model is validated and applied to a case study for the Stockholm metro network. The findings suggest that an increase in peak hour demand leads to a more even passenger distribution among individual train cars, which partially counteracts the increased disutility caused by the higher passenger volumes. Interestingly, the closure of the most popular entrance point at one of the stations leads to lower train crowding unevenness and consequently reduces the on-board travel discomfort. We find that the user’s cost is significantly underestimated when passenger distribution among cars is not accounted for.

Keywords Public transport · Transit assignment · Crowding · Agent-based simulation · Passenger distribution

S. Peftitsi
Division of Transport Planning, KTH Royal Institute of Technology, Stockholm, Sweden
E-mail: soumela@kth.se

E. Jenelius
Division of Transport Planning, KTH Royal Institute of Technology, Stockholm, Sweden

O. Cats
Department of Transport and Planning, Delft University of Technology, Delft, The Netherlands & Division of Transport Planning, KTH Royal Institute of Technology, Stockholm, Sweden
1 Introduction

Many public transport systems are subject to overcrowding during peak periods. On-board crowding is associated with many negative effects on passengers, such as increased discomfort and stress, delays and denied boarding (Tirachini et al., 2013). There is a need to improve system performance and the level of service. A thorough understanding of public transport users’ travel behavior under crowded conditions as well as their motivation when making boarding decisions is thus essential for increasing the comfort and efficiency of public transport.

Passenger loads can be highly unevenly distributed along platforms and in individual cars of trains and metros even during peak hours (TRB, 2014; Zhang et al., 2017). Kim et al. (2014) investigated the causes behind the uneven passenger distribution between train cars, concluding that passengers’ motivation for minimizing the walking distance at the destination is the most decisive factor for choosing a specific train car. Some studies aim to reduce the unevenness of the on-board passenger distribution by determining the optimal train stop location along a platform (Sohn, 2013) or through real-time crowding information system (Zhang et al., 2017).

Transit assignment models (TAM) are used for describing how passengers are distributed in the public transport network, predicting passengers’ travel route decisions and modeling crowding. However, these models do not evaluate crowding in individual train cars. These models implicitly assume that train cars are evenly crowded, hence yielding an unrealistic vehicle capacity utilization and result with an underestimation of user’s cost.

The objective of this study, motivated by the aforementioned shortcoming of the state-of-the-art TAM, is to propose a quantitative methodology for analyzing capacity utilization of individual train cars. The main contributions of the paper are:

– The development of a quantitative approach for describing passengers’ car-specific boarding choices and evaluating crowding in individual train cars.
– Incorporating the passenger path choice modeling in a dynamic and stochastic TAM which captures train car choices.
– The developed model accounts for day-to-day learning, where passengers’ decisions are made in a repetitive way, taking the impact of car-specific crowding into account. This allows to embed the ‘local’ train car loadings into the ‘global’ passenger route choice.
– The validity of the model is investigated by comparing the simulated output against an empirical data set.
– The effect of demand and infrastructure changes on crowding distribution among individual train cars is investigated using simulation scenarios.

In the proposed model, transit stops are divided into sections, each corresponding to a specific train car. Each car is modelled as a separate unit with a corresponding occupancy and capacity. Using an agent-based simulation model, the distribution of passengers between individual cars is dynami-
Evaluating train crowding unevenness 3
cally and stochastically modelled. The rapid rise of information and communication technologies enables the collection of massive and continuous mobility data that facilitate calibration and validation of car-specific transit assignment models.

The remainder of the paper is structured as follows. In section 2 we review the relevant literature. The dynamic model for evaluating the distribution of crowding across a multi-car vehicle is presented in Section 3. Next, we present in section 4 the application of the proposed model to a case study in Stockholm, followed by the presentation of results in Section 5. Section 6 concludes with a discussion of the key findings, reflecting on the limitations of the study and providing an outline of directions for further research.

2 Literature review

Public transport users make travel decisions considering a variety of factors, namely travel time, comfort level and station physical layout characteristics. To avoid on-board crowding passengers make trade-offs, i.e. choose an alternative travel path, board a less crowded train car, wait for the next train service or adapt their departure time to ensure a seat (Pownall et al., 2008; Kim et al., 2015). The skewness of passengers distribution along the station platform and eventually among individual cars of the vehicle is closely related to the physical layout of the station and the location of entrance and exit points (Szplett and Wirasinghe, 1984; Krstanoski, 2014; Liu et al., 2016). Data-driven approaches for studying the behavioral aspects of passengers’ boarding choices find that passengers choose a specific metro car to board aiming to minimize the walking distance and on-board discomfort (Kim et al., 2014; Peftitsi et al., 2020).

A variety of models exist which describe the assignment of passengers to a transit network, predict travelers’ decisions and model crowding at station and inside transit vehicles. The growing literature on TAM which are broadly classified into frequency-based and schedule-based approaches to modeling transit path choice, is reviewed by Fu et al. (2012). Frequency-based models represent transit services at the line level. Spiess and Florian (1989) proposed the first frequency-based assignment framework based on passengers’ optimal strategies for minimizing the generalized travel time. Frequency-based approaches for congested networks were presented by Lam et al. (1999), accounting for the total passenger travel cost and by Cepeda et al. (2006), accounting for service capacity. Schmöcker et al. (2008) introduced a dynamic frequency-based transit assignment model, considering passengers that fail to board due to insufficient remaining capacity.

Schedule-based assignment models represent individual train trips that their availability is dictated by timetables. A schedule-based model, presented by Nuzzolo et al. (2001), accounts for on-board congestion by using a passenger discomfort factor. Vehicle capacity constraints were introduced by Nguyen et al. (2001); Papola et al. (2009), to model boarding passengers considering
the residual capacity of the vehicle and the possibility of denied boardings. Poon et al. (2004) used a schedule-based traffic assignment model for congested transit networks with capacity constraints, to predict the queuing time per passenger, assuming that passengers are queuing according to the first-in-first-out (FIFO) rule. Another model that takes into account transit schedules and vehicle capacity to assign passengers to paths and model the impact of priority rules was proposed by Hamdouch and Lawphongpanich (2008). On-board passengers have priority and the waiting passengers are assumed to board the vehicle in a FIFO or at a random manner. Seat capacity constraints have been considered to model the on-board discomfort effect on the sitting and standing passengers (Leurent, 2010; Sumalee et al., 2009).

Agent-based simulation models mimic individual passengers’ behavior and choices and allow thus to model dynamic congestion effects. Wahba and Shalaby (2005) were among the first to propose a framework based on the assumption that individual passengers adjust their travel behavior based on their experience. Zhang et al. (2008) developed an agent-based simulation model to capture passenger boarding and alighting movements at stops. Häseler et al. (2020) presented a model for describing the interaction between passenger movements on platform and inside the vehicle. Rexfelt et al. (2014) have focused on modeling the behavior of individual passengers at stops and on-board buses, assessing the effect of vehicle layout on boarding and alighting passenger movements. A dynamic and stochastic TAM which captures congestion and crowding effects (denied boarding, on-board crowding and service irregularity), was proposed by Cats et al. (2016). Cats and Hartl (2016) compared the ability of schedule-based and agent-based TAM to model on-board congestion, finding that the latter is more sensitive to variations in demand.

Although tools for modeling and predicting passenger flows in public transport networks are widely used, there is limited knowledge on how to model crowding distribution among individual train cars. This motivates the need for developing a dynamic model to capture passengers’ boarding choices of individual train cars and evaluate the effects of crowding unevenness.

3 Methodology

Modelling the distribution of passengers on-board multi-car urban rail vehicles involves changes to both supply and demand representation and processes. From the supply side, transit vehicles and station platforms representation requires the identification of car units and platform sections, respectively, to enable modelling the distribution of passengers along the vehicles. From the demand side, path choice modelling in an existing public transit simulation model requires extension to capture passengers’ car boarding choices. The developed model takes the effect of car-specific crowding through a day-to-day learning process into account.
3.1 Simulation modeling approach

A dynamic agent-based public transport operations simulation model, BusMezzo, is used in this study for modeling the congestion effects on-board the vehicle, considering vehicle capacity constraints (Cats, 2013). Crowding is evaluated at the vehicle level and the distribution of passengers over vehicle units (e.g., rail cars) is not captured. The model simulates individual passenger path decisions and the movements of individual transit vehicles. The transit system is dynamically and stochastically represented and integrated into Mezzo, which is an event-based mesoscopic traffic simulation tool (Toledo et al., 2010). Different public transport modes, i.e., metro, commuter train, bus and tram, are modeled using different vehicle types with distinct capacity characteristics and dwell time functions. A set of trips is assigned to each vehicle type and hence, BusMezzo models also the propagation of delays caused by trip chaining.

The transit network in BusMezzo includes a set of transit stops $S$ and a set of transit lines $L$. Each transit stop $s \in S$ may be served by more than one transit line. In this study, we extend model functionality to allow for modeling passenger distribution over vehicles. To enable modeling passenger distribution over vehicles, each transit stop $s$ is divided into $K_s$ sections. Each transit line $l$, defined by an Origin-Destination (OD) pair and a sequence of stops $S_l$, is served by a set of trips denoted by $J_l$. Each vehicle serving trip $j \in J_l$ consists of $I_j$ cars. It is assumed that each transit stop is operated by transit lines with the same number of car units per vehicle (i.e., $I_j = K_s$ for all $s \in S_l$ and $j \in J_l$) and hence, each platform section corresponds to a certain car unit.

3.2 Passenger path choice modeling

Passengers are generated stochastically according to Poisson processes based on OD matrices. Each origin and destination is specified as a pair of platform sections located at certain stations. Throughout the simulation, each passenger makes a sequence of path decisions, namely boarding, alighting and walking decisions, that combined yield the realization of a path. The path decisions are described with random utility discrete choice models. Each alternative is associated with a utility, evaluated based on the passenger’s preferences and expectations, which are shaped by prior knowledge, gained experience and available provided information (Cats and Gkioulou, 2017).

In BusMezzo, each transit path alternative $a$, which connects an origin location $o$ to a destination $d$ and is included in path set $A^{od}$, is defined as a combination of stops, lines and walking links (Cats et al., 2016). To capture crowding in individual cars of a multi-car vehicle, the path alternative is further defined in this study as an ordered combination of transit stops associated with a platform section, transit lines associated with a car unit and a set of walking links between stops as well as platform sections.
Each feasible path alternative \( a \) is associated with a utility; the deterministic part of the utility for passenger \( y \) of a feasible path \( a \) is:

\[
v_{y,a} = \beta_{a}^{inv} t^{inv}_{y,a} + \beta_{a}^{wait} t^{wait}_{y,a} + \beta_{a}^{walk} t^{walk}_{a} \quad \forall y \in Y, a \in A^{od}
\]

where \( t^{wait}_{y,a} \) is time-dependent and passenger-specific waiting time that depends on passenger arrival process and service frequency, \( t^{inv}_{y,a} \) is the expected total in-vehicle time, \( t^{walk}_{a} \) is the expected total walking time included in the path alternative and \( \beta \)'s are the corresponding utility function coefficients.

In order to obtain passenger loads on individual cars, we need to incorporate the train car choice into the dynamic individual path choice making. In the following we detail how the path choice model accounts for train car choice and related travel attributes. Any decision in the simulation model involves a passenger making a decision regarding the following element, i.e stop associated with section, line associated with car unit and walking link, considering all the path alternatives that are associated with the specific element. When passengers reach the end of a path element, they choose the next path element that maximizes their expected utility. The utility that passenger \( y \) associates with a path element \( c \) (\( c \in C \)), denoted by \( u_{y,c} \), is given as composite utility of all path alternatives \( A^{od} \):

\[
u_{y,c} = \ln \sum_{a \in A^{od}} e^{v_{y,a}}
\]

The probability that a passenger \( y \) will choose the next path element \( c \) is then

\[
P_{y,c} = \frac{e^{u_{y,c}}}{\sum_{c \in C} e^{u_{y,c}}}
\]

Passenger path choices involve three types of decisions as described in the following.

**Walking decision:** The passenger path choice process starts with the walking decision. Passenger \( y \in Y \) decides whether to stay at the origin stop \( o \) or to walk to some platform section \( k \) of some transit stop \( s \). Each time a passenger alights from a transit vehicle, a new origin location is set and another walking decision needs to be made. The walking utility is based on the walking distance to a platform section of the first stop that the passenger wants to walk to and on the expected future travel attributes, including in-vehicle, walking and waiting times for all path alternatives between this section and traveler’s final destination. The total walking distance of the downstream walking links includes on-platform walking distances at the destination stop and it is also included in the walking decision. This allows capturing travellers’ choice to minimize walking time at the downstream station while waiting, depending on the location of the desired exit. In the current model implementation, section-to-section walking distances are computed based on the shortest section-to-section walking path between two stops.

**Boarding decision:** Each time a transit vehicle \( j \) arrives at transit stop \( s \), passenger \( y \) makes a boarding decision; board the vehicle or stay on the
platform. In the boarding decision process, the utility associated with boarding is compared to the utility associated with staying and waiting for other vehicle. Expected in-vehicle, walking and waiting times are involved in the boarding utility function. The number of passengers boarding car $i$ of vehicle run $j$ at stop $s$, denoted by $q_{ij}^{board}$, is given by the number of passengers that make a positive boarding decision if the car has not reached its capacity $\gamma_i$; otherwise, the number of boarding passengers is equal to the available capacity of the car. If the car, that is adjacent to the waiting platform section $s_k$, has reached its capacity, the passenger stays on the same platform section waiting for other vehicles included in the choice set.

**Alighting decision**: Upon boarding a transit vehicle $j$, passenger $y$ decides at which downstream transit stop to alight. The platform section $k$, that the passenger will alight at, is already determined by the car unit $i$ that the passenger has boarded, under the assumption that passengers do not move between car units (this may also not be possible for some vehicle configurations). The number of passengers alighting from car $i$ at platform section $s_k$ equals to the total number of passengers that make a positive alighting decision.

The number of passengers on-board car $i$ of vehicle run $j$ when the vehicle departs from a stop $s$, denoted by $q_{ij}^{onboard}$ is a function of alighting and boarding flows into and from car $i$. In BusMezzo passengers are allocated to seats assuming a First-In-First-Out (FIFO) rule, where seats of each car are filled sequentially and standing passengers exist when the car seated capacity has been reached.

Figure 1 illustrates the path alternative definition for passengers that start their trip at origin location $o$ and aim to reach destination location $d$. For illustration, it is assumed that the transit stops are operated by 3-car trains and hence, the station platforms are divided into three sections. The passenger, starting at $o$, has three alternative connection choices that can be accessed by walking; the first, second and third platform sections of transit stop $s_1$, denoted by $s_{1,1}$, $s_{1,2}$ and $s_{1,3}$, respectively. The stop is served by the transit line $l_1$, while each platform section is served by the corresponding train car unit of the line, denoted by $i_{1,1}$, $i_{1,2}$ and $i_{1,3}$ for the first, second and third car, respectively. A transit user that decides to make a walking connection to the first section of the stop $s_{1,1}$ will board $i_{1,1}$, if they make a boarding decision, considering car capacity constraints, and will alight at the first platform section of the transfer transit stop $s_{2,1}$, which is then set as a new origin transit location. From the alighting platform section, the passenger makes a new walking decision to a platform section of the same or different stop. For the origin-destination pair illustrated in Figure 1, nine path alternatives are available.

*Perceived in-vehicle travel time*

The developed model accounts for day-to-day dynamics, where an iterative network loading is performed to mimic passengers’ adaptive travel behavior in the real world. Passengers’ travel decisions are made in a repetitive way and their expectations about car-specific crowding, are updated on a day-to-day
basis. Passengers store information based on the experience gained during the previous days and they may alter their travel strategy accordingly.

When day-to-day dynamics are not considered, i.e. the first simulated day, passengers choose a section of the first connected stop and thereby a train car based solely on the walking distance to the platform section and on the expected utilities of the path alternatives available between this section and the desired exit at the destination. In this case, passengers expect equally utilized cars of the next arriving train. Performing an iterative network loading, car boarding choice is also affected by passenger’s expectations about car crowding. In this case, in-vehicle travel times, included in the utility associated with a path alternative, are weighted with a car-specific crowding factor based on passenger’s gained experience. Consequently, passengers expect different crowding levels in individual cars.

3.3 Data requirements

Advances in information and communication technologies in public transport enable the generation of big data sources, such as cell phone data, social media data or smart-card data, that can be utilized for modeling travel behavior and predicting passengers’ movements as well as investigating travel patterns (Chen et al., 2016).

For model application, the platform section-level OD information is required to represent the passenger demand for each pair of platform sections of a given OD pair. If this information is not readily available, three types of data are required to represent the demand.

Average station-to-station travel demand data for each OD pair describe the average number of trips between a given origin and destination. Such data may be obtained through cell phone data (Alexander et al., 2015; Toole et al., 2015; Bachir et al., 2019) or automated fare collection (AFC) data. AFC data play an increasingly important role in estimating travel demand in public transport systems (Munizaga and Palma, 2012; Alsger et al., 2016).
Pedestrian incoming and outgoing flows at each access point of the station are useful to describe the passenger movements at the entrance level and can be used to estimate the probability that a passenger initiates or ends the trip at a certain section of the platform. This information may be obtained through passenger counts, AFC data that has been utilized in Ingvardson et al. (2018) to model passengers’ arrival patterns and in Peftitsi et al. (2020) to analyze how passengers make metro car boarding choices or cell phone data that has been used in Aguiléra et al. (2014) to measure passenger flows in a public transit system.

The physical infrastructure characteristics of the network, including the dimensions of the platform and location of entrance and exit points, are also required to define the stop characteristics and walking distances within stations as well as between stops.

For model validation, passenger load data for each car unit, describing the crowding level on-board individual cars, are required. Car load data may be collected from car weight measurements or obtained through sensors installed at the car doors.

4 Application

4.1 Case study description

The proposed modeling framework is applied to a case study for the metro network in Stockholm. The Stockholm metro system is used by more than 1 million passengers per workday. Although passenger loads are close to capacity during peak hours, passengers are often unevenly distributed among train cars and 20% of the seats remain unoccupied during the morning peak hour (SL, 2017). The model is applied to the southbound segment of metro line 14 between Mörby centrum (MÖR) and Stadion (STD), which operates with a planned headway of 5 minutes during the morning peak period (06:00 - 09:00 am). The segment exhibits high average on-board passenger load. The studied area is shown in Figure 2.

The passenger distribution on-board the trains is highly skewed towards the front car during the morning peak hour (Figure 3). On average, 41% of the on-board passengers occupy the front train car, while the rear car is occupied by only 25% of the passengers.

4.2 Network representation

The transit network representation in BusMezzo includes both directions of the metro line 14 in Stockholm. The network consists of 39 stops which are served by 72 vehicle trips, each of which consists of 3 train cars. The transit network is represented in BusMezzo with detailed vehicle scheduling, the connectivity of stops by walking as well as the shortest section-to-section walking path.
between these stops, required to compute the section-to-section distance of the walking links.

4.3 Demand representation

Passenger demand is simulated for the morning peak hour. The OD travel demand data for the morning peak period at the station-to-station level are taken from the transit assignment model Visum, based on the official planning zonal OD matrix. For each OD pair, a platform section demand matrix indicating the probability that a passenger starts and ends the trip at a certain platform section at the origin and destination stop, respectively, was produced based on the total incoming and outgoing passenger flows at each entrance point of the station, which are obtained through passenger counts in the morning peak hour.

4.4 Scenarios design

We simulate the following three scenarios in BusMezzo:

1. *Base scenario*, where the case study is simulated with the current average morning peak hour demand.
2. **Increased demand scenario**, where the case study is simulated with the current average morning peak hour demand increased by 30%.

3. **Intervention scenario**, where an infrastructure change is considered at one of the metro stations, namely Danderyds sjukhus (DAS), which is a station with two access points located at the south and north ends. The south, which is the most popular access point, is considered as temporarily non-available in this scenario. Elevator and escalator maintenance is one of the factors that might require the temporary closure of a station entrance point.

Since the transit simulation model BusMezzo is stochastic, each scenario needs to be evaluated based on a number of simulation replications. The number of replications required $N(m)$, given $m$ initial runs, as it is given in Cats et al. (2010), is determined by

$$N(m) = \left( \frac{\sigma(m)t_{m-1,1-\alpha/2}}{\mu(m)\varepsilon} \right)^2$$  \hspace{1cm} (4)

where $\sigma(m)$ is the standard deviation of the average generalized travel cost per passenger of $m$ simulation runs, $t_{m-1,1-\alpha/2}$ is the critical value of the t-test for $m-1$ degrees of freedom and level of significance $\alpha$, $\mu(m)$ is the mean generalized travel cost per passenger of $m$ simulation runs and $\varepsilon$ is the allowed error.

Given significance level and allowed error of 5%, 10 simulation runs were found sufficient, yielding a maximum error of 5%.
4.5 Performance evaluation

The impact of alternative scenarios on the performance of the system is evaluated by considering the average on-board crowding unevenness across the vehicle, the average boarding passengers unevenness and the average generalized travel cost per passenger.

4.5.1 Crowding unevenness

Having a single metric for measuring crowding unevenness facilitates comparisons between different passenger distributions. The distribution of passengers among the cars of a train run \( j \) upon departure from stop \( s \) is systematically measured using the Gini coefficient \( G_{js} \).

\[
G_{js} = \frac{1}{2|I|} \sum_{i=1}^{I} \sum_{i'=1}^{I} |q_{onboard}^{ij} - q_{onboard}^{i'j}|
\]

(5)

This train crowding unevenness metric measures how much the passenger load distribution deviates from the totally equal distribution, i.e. when all train cars are equally utilized. The metric takes the value 0 in case of perfect evenness in the train - on-board crowding is minimal given the overall passenger load level, i.e. passengers are equally distributed over all train cars - and the value 1 in case of perfect unevenness - on-board crowding is maximal given the overall passenger load level, i.e. passengers are filling cars in succession.

On-board train crowding distribution is based on passengers’ boarding behavior at the station and hence it is essential to evaluate the performance of the system by considering the distribution of boarding passengers. Similarly, the distribution of the passengers boarding the train \( j \) at stop \( s \) is given by

\[
G_{board}^{onboard} = \frac{1}{2|I|} \sum_{i=1}^{I} \sum_{i'=1}^{I} |q_{board}^{ij} - q_{board}^{i'j}|
\]

(6)

4.5.2 Generalized travel cost

The generalized travel cost is defined as the weighted sum of in-vehicle, walking and waiting time.

The disutility of in-vehicle time reflects on-board passenger discomfort. In-vehicle travel time is weighted with an on-board car-specific crowding factor, that is defined as a non-linear function of the load factor of car \( i \), and the corresponding user-specified parameter for in-vehicle time \( \beta_{inv} \). Crowding factor value varies between sitting and standing passengers (Wardman and Whelan, 2011).

Walking and waiting times are weighted with the corresponding user-specified parameters for walking \( \beta_{walk} \) and waiting time \( \beta_{wait} \), respectively. Time valuations are based on the valuations reported in the literature (Wardman, 2004). The value of in-vehicle, walking and waiting time are set to: \( \beta_{inv} = -1, \beta_{walk} = \beta_{wait} = 2 \cdot \beta_{inv} = -2 \).
5 Results

We first study the performance of the model in the studied metro line segment to assess the validity of the model, and then investigate the effect of demand and infrastructure changes on crowding unevenness using simulation scenarios.

5.1 Model validation

The simulation outputs are tested against an empirical data set of car load data to investigate the validity of the simulation model (Figure 4). Empirical car load data estimated through the car weight measurements, based on an average of 78kg per passenger including luggage, are available at train departure from each stop on the studied line segment for the morning peak period in October 2016. The outputs of a simulation hour are tested for the highest peak hour of the morning rush period (07:30 - 08:30 am). The relation between in-vehicle and walking time valuation is calibrated to investigate the sensitivity of the model and assess the goodness of fit. We conduct several t-tests to determine if there is significant difference between the mean car passenger loads of the simulated and empirical data set at train departure from the studied stops when the model accounts for day-to-day learning and crowding effect is taken into account in decision making process. We find that the model best reproduces the empirical car passenger loads when passengers value walking time higher. Car loads upon train departure from stops between BEH and STD, as well as the passenger loads of the most crowded cars at MOR and DAS can be well reproduced. The lower accuracy of predictions for the first two stops at the case study corridor might be explained by the more diverse travel strategies due to a low demand level.
The impact of modelling car-specific perceived in-vehicle travel times on train crowding unevenness, based on the Gini coefficient, is illustrated in Figure 5. Slightly lower Gini coefficient values, indicating more even average passenger distributions across the train, are observed when car-specific perceived in-vehicle travel times are taken into account in the iterative network loading process. We find that the average perceived in-vehicle time per passenger is 0.4% shorter when car-specific perceived in-vehicle travel times are accounted for in passengers’ decisions since they yield lower crowding unevenness (Figure 6). Experienced passengers are expected to alter their travel behavior aiming to minimize car-specific discomfort when the train passenger volumes are higher. In increased demand conditions, experienced passengers walk more, making trade-offs between on-platform walking and on-board crowding level (Figures 7 and 8). In particular, average walking time increases by 3%, leading to passengers experiencing 2% lower on-board discomfort due to the more even train crowding distribution.

5.2 Model application

5.2.1 Crowding unevenness

Figure 9 shows the average unevenness of boarding passengers, for the three simulated scenarios: the Base scenario, the Increased demand scenario and the Intervention scenario (Section 4.4). In all simulated scenarios, day-to-day learning is used to incorporate car-specific in-vehicle travel times in passengers’ train car choices. On average, boarding passenger unevenness drops by
Fig. 6 Average generalized travel time per passenger of the simulated output (with and without day-to-day dynamics).

Fig. 7 Average on-board train crowding unevenness, given by the Gini coefficient, of the simulated output (with and without day-to-day dynamics) in increased demand conditions (130%).

0.8 percentage points under increased demand conditions. Considering the large passenger volumes in some of the cars of the arriving train, experienced boarding passengers are skewed towards the next closest cars, leading to a more even passenger distribution. This stems from the increasing discomfort associated with increased crowding as implied by the established values of in-vehicle time multipliers in the literature. The intervention scenario has a significant impact on the distribution of boarding passengers at DAS, leading
Fig. 8 Average generalized travel time per passenger of the simulated output (with and without day-to-day dynamics) in increased demand conditions (130%).

Fig. 9 Average boarding passenger unevenness for base, increased demand and intervention scenarios.

...to high unevenness, since passengers’ preference has switched to the car located close to the single access point. However, passengers’ car choice at most of the downstream stations is not significantly affected by the intervention at the upstream station.

Figure 10 presents the average crowding unevenness on-board trains upon departure from each stop. We find that under increased demand conditions the average train crowding unevenness decreases slightly by 0.4 percentage points at the studied stations, explained by passengers’ car boarding choices...
shown in figure 9. Crowding unevenness is large on-board trains departing from DAS due to the increased passengers’ preference for the rear train car caused by the closure of the southern station entrance point, but it gradually decreases at the downstream stops (Figure 10). The infrastructure intervention has the most significant impact at the two most crowded stops, TEH and STD, where train crowding unevenness halved with a drop of approximately 6.5 percentage points, leading to a much more even passenger load distribution (Figure 11). The impact of the infrastructure intervention arguably depends on the entrance point which is disrupted as well as the unevenness of crowding at the downstream stations.

5.2.2 Generalized travel cost

The average generalized time components per passenger are used to evaluate the effects of the simulated scenarios on user cost (Figure 12). In the increased demand scenario, the average perceived in-vehicle time per passenger increases by 19%. This finding suggests that the more even passenger distribution among individual train cars, partially counteracts the increased disutility caused by higher passenger volumes. In increased demand conditions, experienced passengers walk more to attain lower on-board discomfort. The highly skewed demand distribution towards the rear train car at DAS, due to the closure of the most popular entrance point, cancels out the crowding unevenness at the downstream stations. As expected, passengers starting the trip at UNT and TEH experience lower crowding discomfort and consequently the average perceived in-vehicle time per passenger decreases by 18% and 5%, respectively, when compared to the base scenario.
Fig. 11 Average car passenger load for base, increased demand and intervention scenarios.

If individual train cars were not modelled, the average generalized travel time per passenger would be 20% shorter for the Base scenario, indicating that the user’s cost is significantly underestimated, when passenger distribution among cars is not accounted for.

6 Conclusions

We propose a modelling framework that evaluates and forecasts crowding distribution on-board multi-car vehicles. Each car is modelled as a separate unit...
with its own capacity constraints, which enables to capture passenger flows and crowding on-board each car unit. The path choice modelling implemented in transit assignment simulation model has been extended to allow for the inclusion of platform section and car unit in the walking and boarding decision process, respectively.

An application of the proposed modelling framework to a six station segment of the southbound direction of metro line 14 in Stockholm, where on-board crowding is on average highly skewed towards the front cars of the metro trains, is conducted. The developed model accounts for day-to-day dynamics, where platform section and eventually car choices are affected by passengers' expectations about car-specific on-board crowding based on the experience they gain during the course of successive network loadings. The validity of the model and its sensitivity to time valuations as well as day-to-day learning has been examined. We conclude that the model can reproduce the average empirical car passenger load data when crowding effect is taken into account in the iterative network loading model and walking time induces a higher disutility. This finding suggests that passengers may dislike walking along the platform more than walking from/to the station.

Increased demand level reduces crowding unevenness on-board trains upon departure from the studied stops. This finding suggests that when passenger load is closer to capacity, i.e. limited residual car capacity, there is lower flexibility in relation to how passengers are distributed among the cars of the train. Although passengers are more evenly distributed among individual train cars when travel demand increases, the perception of in-vehicle travel time increases, due to more severe on-board car crowding conditions. The closure of a popular station entrance point due to i.e. maintenance work required, leads to highly uneven distribution of boarding passengers at the specific stop, which is skewed towards the single access point, but it cancels out the crowding unevenness on-board when departing from the subsequent stops. We find that for the same demand level, passengers' perceived in-vehicle time is lower when crowding is more evenly distributed across the train.

The developed model can be used for decision support by public transport authorities and operators to evaluate the effects of possible infrastructure investments or operational changes on crowding on-board trains. Crowding effects are evaluated in a more realistic manner, by considering that passengers are not evenly loaded among individual train cars. Emerging data sources, such as pedestrian counts and pedestrian density measurements available from cameras and sensors, can be in the future utilized to further calibrate and validate the developed car-specific transit assignment model.

In the developed simulation model, dwell times are modelled as a monotonically increasing function of the total number of boarding, alighting and on-board passengers. Car-specific crowding level is not included in dwell times modelling, which may result in underestimated dwell times. The model could be extended to study how car capacity utilization may affect train dwell times. There is lack of behavioral knowledge on the trade-off between walking and waiting when those are not exact substitutes, i.e. whether passengers choose to
walk while waiting on the platform at the origin station as opposed to walking on the platform at the destination station. Future research on the disutility of walking when it is substituting waiting will thus allow for refining model specification.

Future research should also include the modelling of passenger movements along the platform in order to evaluate on-platform crowding unevenness and its interaction with on-board crowding unevenness. Another direction for future research is to evaluate the impact of real-time crowding information systems on the distribution of passengers across a multi-car vehicle. Such a passenger information system has been tested in a pilot study and it was found that real-time crowding information system has a statistically significant positive impact on the car choice (Zhang et al., 2017). From a simulation perspective, Drabicki et al. (2017) formulated a path choice model accounting for passengers’ access to crowding information that is consistent with the agent-based modeling approach adopted in this study.

Acknowledgements This work was funded by Stockholm County Council (SLL), Centre for Transport Studies (CTS) and TRENop Strategic Research Area. The authors would like to thank MTR Nordic for providing the metro train car load data.

References


Evaluating train crowding unevenness


